## **RAW SEQUENCE LISTING**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number:	10/534.099A
Source:	PU,
Date Processed by STIC:	7//4/06
•	

## ENTERED



PCT

RAW SEQUENCE LISTING DATE: 07/14/2006
PATENT APPLICATION: US/10/534,099A TIME: 10:18:51

Input Set : F:\UTSC772US.APP.txt

Output Set: N:\CRF4\07142006\J534099A.raw

```
3 <110> APPLICANT: DE CROMBRUGGHE, BENOIT
             AKIYAMA, HARUHIKO
      6 <120> TITLE OF INVENTION: HA4, A NEW OSTEOBLAST- AND CHONDROCYTE-SPECIFIC SMALL
             SECRETED PEPTIDE, COMPOSITIONS AND METHODS OF USE
      9 <130> FILE REFERENCE: UTSC:772US
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/534,099A
     12 <141> CURRENT FILING DATE: 2005-05-06
     14 <150> PRIOR APPLICATION NUMBER: PCT/US2003/035139
     15 <151> PRIOR FILING DATE: 2003-11-04
     17 <150> PRIOR APPLICATION NUMBER: 60/423,690
     18 <151> PRIOR FILING DATE: 2002-11-04
     21 <160> NUMBER OF SEQ ID NOS: 3
     23 <170> SOFTWARE: PatentIn Ver. 2.1
     25 <210> SEQ ID NO: 1
     26 <211> LENGTH: 735
     27 <212> TYPE: DNA
     28 <213 > ORGANISM: Artificial Sequence
     30 <220> FEATURE:
     31 <221> NAME/KEY: CDS
     32 <222> LOCATION: (1)..(735)
     34 <220> FEATURE:
     35 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
     36
             Primer
     38 <400> SEQUENCE: 1
     39 atg cac ccc caa ggc cgc gcg gcc ccc ccg cag ctg ctc ggt ctc
     40 Met His Pro Gln Gly Arg Ala Ala Pro Pro Gln Leu Leu Gly Leu
     41
                                                                  15
     43 ttc ctt gtg ctg ctg ctt cag ttg tcc gca ccg tcc agc gcc tct
                                                                           96
     44 Phe Leu Val Leu Leu Leu Gln Leu Ser Ala Pro Ser Ser Ala Ser
                     20
                                         25
     47 gag aac ccc aag gtg aag caa aaa gcg ctg atc cgg cag agg gag gtg
                                                                           144
     48 Glu Asn Pro Lys Val Lys Gln Lys Ala Leu Ile Arg Gln Arg Glu Val
                 35
                                     40
     51 gta gac ctg tat aat gga atg tgt cta caa gga cca gca gga gtt ccc
                                                                           192
     52 Val Asp Leu Tyr Asn Gly Met Cys Leu Gln Gly Pro Ala Gly Val Pro
     53
             50
                                 55
     55 ggt cgt gat ggg agc cct ggg gcc aat ggc att cct ggc aca cct ggc
                                                                           240
     56 Gly Arg Asp Gly Ser Pro Gly Ala Asn Gly Ile Pro Gly Thr Pro Gly
     59 atc cca ggt cgg gat gga ttc aaa ggg gaa aag gga gaa tgc tta agg
                                                                           288
     60 Ile Pro Gly Arg Asp Gly Phe Lys Gly Glu Lys Gly Glu Cys Leu Arg
```

63 gaa agc ttt gag gag tcc tgg acc cca aac tat aag cag tgt tcg tgg

336

Input Set : F:\UTSC772US.APP.txt

Output Set: N:\CRF4\07142006\J534099A.raw

```
64 Glu Ser Phe Glu Glu Ser Trp Thr Pro Asn Tyr Lys Gln Cys Ser Trp
               100
                                   105
                                                                      384
67 agt tcg ctg aac tat ggc ata gat ctt ggg aaa att gcg gag tgt aca
68 Ser Ser Leu Asn Tyr Gly Ile Asp Leu Gly Lys Ile Ala Glu Cys Thr
                               120
71 ttc acg aag atg cgc tcc aac agt gct ctg cga gtt ctg ttc agt ggc
72 Phe Thr Lys Met Arg Ser Asn Ser Ala Leu Arg Val Leu Phe Ser Gly
73
       130
                           135
75 tca ctt cgg ctc aaa tgc agg aat gca tgc tgt cag cgc tgg tat ttt
                                                                      480
76 Ser Leu Arg Leu Lys Cys Arg Asn Ala Cys Cys Gln Arg Trp Tyr Phe
                                           155
                       150
79 aca ttt aat gga gct gaa tgt tca gga cct ctt ccc atc gaa gcc atc
                                                                      528
80 Thr Phe Asn Gly Ala Glu Cys Ser Gly Pro Leu Pro Ile Glu Ala Ile
                                       170
                  165
                                                                      576
83 atc tat ctg gac caa gga agc cct gag tta aat tca act att aat att
84 Ile Tyr Leu Asp Gln Gly Ser Pro Glu Leu Asn Ser Thr Ile Asn Ile
               180
                                                                      624
87 cat cgt act tcc tct gtg gaa gga ctc tgt gaa ggg att ggt gct gga
88 His Arg Thr Ser Ser Val Glu Gly Leu Cys Glu Gly Ile Gly Ala Gly
                               200
91 ttg gta gat gtg gcc atc tgg gtt ggc acc tgt tca gat tac ccc aaa
                                                                      672
92 Leu Val Asp Val Ala Ile Trp Val Gly Thr Cys Ser Asp Tyr Pro Lys
                           215
95 gga gac gct tct act gga tgg aat tcc gtg tct cgc atc atc att gaa
                                                                      720
96 Gly Asp Ala Ser Thr Gly Trp Asn Ser Val Ser Arg Ile Ile Ile Glu
97 225
                       230
                                           235
99 gaa cta ccg aaa taa
                                                                      735
100 Glu Leu Pro Lys
104 <210> SEQ ID NO: 2
105 <211> LENGTH: 244
106 <212> TYPE: PRT
107 <213> ORGANISM: Artificial Sequence
109 <220> FEATURE:
110 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
          Peptide
113 <400> SEQUENCE: 2
114 Met His Pro Gln Gly Arg Ala Ala Pro Pro Gln Leu Leu Gly Leu
116 Phe Leu Val Leu Leu Leu Gln Leu Ser Ala Pro Ser Ser Ala Ser
                 20
118 Glu Asn Pro Lys Val Lys Gln Lys Ala Leu Ile Arg Gln Arg Glu Val
             35
120 Val Asp Leu Tyr Asn Gly Met Cys Leu Gln Gly Pro Ala Gly Val Pro
         50
                             55
122 Gly Arg Asp Gly Ser Pro Gly Ala Asn Gly Ile Pro Gly Thr Pro Gly
                         70
124 Ile Pro Gly Arg Asp Gly Phe Lys Gly Glu Lys Gly Glu Cys Leu Arg
126 Glu Ser Phe Glu Glu Ser Trp Thr Pro Asn Tyr Lys Gln Cys Ser Trp
```

Input Set : F:\UTSC772US.APP.txt

Output Set: N:\CRF4\07142006\J534099A.raw

```
100
                                         105
                                                             110
    127
    128 Ser Ser Leu Asn Tyr Gly Ile Asp Leu Gly Lys Ile Ala Glu Cys Thr
                115
                                     120
    130 Phe Thr Lys Met Arg Ser Asn Ser Ala Leu Arg Val Leu Phe Ser Gly
    131
                                 135
    132 Ser Leu Arg Leu Lys Cys Arg Asn Ala Cys Cys Gln Arg Trp Tyr Phe
                             150
                                                 155
    134 Thr Phe Asn Gly Ala Glu Cys Ser Gly Pro Leu Pro Ile Glu Ala Ile
                                             170
    136 Ile Tyr Leu Asp Gln Gly Ser Pro Glu Leu Asn Ser Thr Ile Asn Ile
                                                             190
                     180
                                         185
    138 His Arg Thr Ser Ser Val Glu Gly Leu Cys Glu Gly Ile Gly Ala Gly
                195
                                     200
    140 Leu Val Asp Val Ala Ile Trp Val Gly Thr Cys Ser Asp Tyr Pro Lys
                                 215
    142 Gly Asp Ala Ser Thr Gly Trp Asn Ser Val Ser Arg Ile Ile Ile Glu
                             230
    143 225
     144 Glu Leu Pro Lys
     148 <210> SEQ ID NO: 3
     149 <211> LENGTH: 20138
     150 <212> TYPE: DNA
     151 <213> ORGANISM: Artificial Sequence
     153 <220> FEATURE:
     154 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
              Primer
     157 <220> FEATURE:
     158 <221> NAME/KEY: modified base
     159 <222> LOCATION: (265)..(18855)
     160 <223 > OTHER INFORMATION: N = A, C, G or T/U
     162 <400> SEQUENCE: 3
     163 gaatteggat ceagacaatg agaaggtaac aatacaaage ttettttggg gaagteeagg 60
     164 tactttattc agttgcaaat tagcaaggct caagctcagt ctctggcctc ctagtgggcc 120
     165 atcttgctct atttttttct ttgtgaaaaa acacagactg aacccctacc ccagattaga 180
     166 accggatctg gaccctttca atcattagtt agtgggtcct tccattttac catggcataa 240
W--> 167 gaattagaag taactggatg ccagnggcaa tccgctgcag actgaccttt aatatcagtt 300
     168 tgcaaaaaat ttagaacaaa taaggcaaag gcaagatgtg cctttggtga ctttggggga 360
     169 tataactgcc cctgttttgt ctaaaaaggc cagaaactct ctggaagtgg aggctgtgct 420
     170 ctaaattaac ttgctaggta attaagaatt ttagctatct acaatatgaa tctctaatat 480
     171 tggactaaat attagaccag tccaagattg aataactggt cacactgaag aaaagagaaa 540
     172 aatcattatg actctattaa atgactaatc ttactaagtc agattatagt ctctgatgtt 600
     173 cttttgccat aaagttaaaa ggcttgaagc tagtccaaac tagaaaaatg gcaagcaaag 660
     174 atgggtctaa aacatgtacc caatttagtt teetagtcae tteaatcagg geactgagte 720
     175 aactcacctg ccagctcatc acacgaatta accatcatgc ttctatagca ttctgtggag 780.
     176 agaacctatg tttttccctt ccctaataag gtagcagttt aggataaaac cataagccaa 840
     177 taagtagatc tttttaagat actataaagc atttattaaa ttctttgaca tctaaatttc 900
     178 aaaattcaaa aataaaagta tgatttaaat agtgtgcacg ggaatgtact aaataatatg 960
     179 cacggggata taatttcccc cttctttgtc ttttaagaag ctaaagtttt aaagtttcac 1020
     180 aatatctagt ccttagaatt aataacttgt tgacaaaaca tctcaaatac ttgactgtat 1080
     181 tcatgtaaca aaggcagtga ctaattgcac tttcaattgt ttttcttaag gagcagttgt 1140
```

Input Set : F:\UTSC772US.APP.txt

Output Set: N:\CRF4\07142006\J534099A.raw

182 aactagaaag cttgaaaagt tattaatagt cacatgtgca taatttattt atttactaga 1200 183 aataaqaqta tttatttqca qcaatttatt aagtaaaggt cctcaagaat taacactatt 1260 184 atgtggaaac aagtagtaac tgaggacact aagaggaaag aattatacat gcacaatttt 1320 185 caagaatact aaaatattat cttaaaccat aactgctttg aatatacaaa gaacaaggtt 1380 186 atatattcaa totatottat atgtacgaco tatagactgo otgagatata aatttagoat 1440 187 ggcattgttt taagatgtcc aggcaatcta gaacaattta ttaaatgtct aaaaaccacc 1500 188 ttcttaaaag gacagcattt tttttctatg agttgtatac gcataagcaa ttgcaaattt 1560 189 gagaattaat agcatctgag ggacgaacaa tcttaagtaa ttgtgagctc tgagatataa 1620 190 cattgactat taatatacaa attaaggttt gattgttcag cattttaaaa caccatctac 1680 191 agcacacaac tagattatct gtgctgaaac ggggggggg gactcaaaag aataagcatg 1740 192 tgatctaatt cacatgtact tacatctcat ataggggttg ttttcaaaca caataattgg 1800 193 gatgtatgca tgtatgaatt tatagaagat acaaaagcat gtataaaaaac aatttttaac 1860 194 ttatctttgg gatattgatt gtcaatctat ttaaaatgat ggtatgtcat aggtcaatta 1920 195 ttaatatttt gtaataacca aataaattgc tgtattgaac aaggaacata cacttagtca 1980 196 ggtttgnttt ttgttttttt gcttttttgn tttgtttntt tttttttttg gnttttcgag 2040 197 acagggtttc tctatagccc ttggctgtcc tggaactcac tttgtagacc aggctggcct 2100 198 cgaactcaga aatccacctg cctctgcctc ccgagtgctg ggatcaaagg cgtgcgccac 2160 199 cacgcccggc tattagtcag gtttttaaga cacaattttt ttttaaaaaa aaaatacacg 2220 200 aatatacatt acaattettt attagtacca cagtggttee aataggatge tgaaactgta 2280 201 ttttqaqqtq aaaaccacct caatcttqqc tqcctqccct cqqccttcct ctcaqqcana 2340 202 agaggetgaa ggeaneecaa tetgttgtag accanaaaac centgeegag tettagtggg 2400 203 aaaaatatgg aggctcatgg ggcaaaggtg aacgcngccc tctcctgagt tcgctgtaaa 2460 204 gccacctgct ccttggcggc ctcttgtacc actatagctc tgggaaactg tgtgtctcaa 2520 205 ggcccaccc cacagcagcc agcagctagc tggcttcaca tgggtcaatg atcggctggg 2580 206 aggatgagaa aaatgacttt acaataatgt tttctcttgc caaagaattg ttttgagcac 2640 207 agctaatatc taattaccat tgattgtaat taataacaat tataaaagct ttctttattt 2700 208 totgtaaago ottotgtagt aaactaaaac ootaagtaat aaaaagatat tgoototgag 2760 209 tcttttgggt gagagcaagg atttaaagta aacttctctt gagaggcatt agctaataaa 2820 210 atattttcca cttaggaaac aatatacact gaaagattaa aactcttggc ttcttgtata 2880 211 gaagcagaaa tgataaaaat ttctcacata atgtagatca atatttagcc atactagagg 2940 212 ctaaatattc cagtngttnc anactcacta gaagcaaaac ctttacaatc attaagangc 3000 213 aaaggagaga gaaaaagaaa caaactttca gatctataat aactatataa aaggtaggca 3060 214 agccagattg taatgccttg ataaattata cagtctgatt gaccctttat aaactctaaa 3120 215 tttgaacttt attaaaacaa catctggcta gatccgtaaa ggtgctcttt tagctaaaag 3180 216 aatateteet ggtttgcaag acaaaggaaa gtcacacaaa cegaacagaa getgetetga 3240 217 gcttagttcc tcttgccatg aggacatata ttagaattac taagtttctt ttgcaccatt 3300 218 aactttggga aagtaaaact ccattttaaa acaatttatc actttcggaa tcaacattaa 3360 219 aactttacta atacattgcg aaatttggcc tgctgtgcca cgtgcttgag aagactcctg 3420 220 agttqccaat tttaagqcta accttctqtt accaatgtaa caattattta atcttaacca 3480 221 ataacttatg agctgattgt gaaaacacgc agagcacatt accaataaaa aaaatgaagc 3540 222 aqttacactt aqaccaatca qaqaatatta atttttctaq ctacaatcat agaataaaag 3600 223 cactttatca tttgccaaat acattaataa cgattgcaga gaaccctcca ggaaaaaacc 3660 224 atttatcage ttttttacce caaaaccaag aggetgeaga etetetttt teetataaag 3720 225 aacagtttct ccagcagggg ccctcctgcg gtgtgctggt ttcctggtta aagcatgtgg 3780 226 ccactcttgg ctttgatgaa ccaccagata aagtttttta gccatcttta ttatccaaca 3840 227 taaaacatag aacattcatt cattctgact ggacacgaac atacatgaat tgaacatgtg 3900 228 aacggattgg tgcacccaaa cacaaaaaca caagagggca gagaacttct gctgtaaggc 3960 229 ccagagagaa caaagcagcg cacccctgaa atttctctct gcttctggga cattttcctc 4020 230 ccacatgatg cagtttctaa ctgtggccat ccacctgatt attaaatccc ttttattaaa 4080

Input Set : F:\UTSC772US.APP.txt

Output Set: N:\CRF4\07142006\J534099A.raw

```
231 ccctttcttt tctcacgcct aaaaaaagca acttcttgga gctgcggcta actgcctgat 4140
232 aaatttcact gctgaaacta agtgaattta gcgcgcaggt ttaacgcagt ttcaacttag 4200
233 cccataccaa ccttctccgg tacgaatttt ctttaattat tttcttcttc catggagctc 4260
234 caaaccaqca atqcttcttc caaatgtcct ctgctgtttc ccagtcccgc gttggttcgc 4320
235 caatctgttg aggccagcca gcggcccaca tgtctgggtt ctagcctggg aggtatcttg 4380
236 qaatctqqaa qaqaaqaqq aactaqqcqq ctcqaqaqaq aatqqaacqa taatqqaacc 4440
237 aagacatcag tetgatcaag gtteaatttt actatttgga gacactgggt tatgaagtag 4500
238 agggagggc ccattectgc caaatcatec ttggagteca gtttcaggtg accaegtgtt 4560
239 ggctccggaa cagctaggcc gcaggtagca gcagtgggag tgacaggctc cacccctgat 4620
240 getetnttag cectaacagt caaacetgat cagecagatt teaggetggg ggggaggtta 4680
242 attqaqacaq qaatttqctq tgtaqccttt attgtcttgg aatttgctct gtaaactagg 4800
243 ctqtccatqa actcacaqaq atctacctgc ctttgcctcc ctagtgctgg gqttaaaggt 4860
244 gttcaccacc accacctagc ccataatctc atattctatc aagggattta ggttccataa 4920
245 atgcacttat cataagaggt tccaacagat ggaacacaac tacacagtat aaggtggact 4980
246 aatacatgtg tgtccttggt gaagaatccc ttatcttatg tcctttcatg tgcttgcttt 5040
247 atcaaaacat cctttcacct gtgtctgctt taggataaca ctccttcaca tgtttgccac 5100
248 agcaaagcac catcagacac gactgacttt ccaaagaacc cttaagtttc cacttcagat 5160
249 aggcatcttc ttctttaggt taagaaaatt tttttcctgt cactctgtgg aaaatatttt 5220
250 ctqtqcattt qacctqtqtt tcttctcctt ccttaattct tattttagat ttggtcactt 5280
251 cacagtgtcc cagatttcct ggatgttttg aaccaggagt tatttttaga ttgaacattt 5340
252 tetttgacet atgtatteat ttettetace atgtetteag tgtetgagat tetettttt 5400
253 aattettgaa ttetgttggt gaagettgee tetgtggttg etgtteaaat teetatattt 5460
254 tcatttccat atttccctca qtttqqqqtc tctttattaa ttctqttgga ggaagggggt 5520
255 ccaqqaqtca cctcacaaat cacatattcc caqqattgat tgatcaqgac caccggccag 5580
256 actcaggage tgaactgtga tgtagagtca agatectaca gggettttaa ageetgagag 5640
257 ctataataac catctctqct aaqttactcc accaatcata acttagggat agggctttct 5700
258 gtaggagcat gtctttgttg tgtacttatt ttgctcctat tggttagggt attcaactat 5760
259 ggcagaggac ttgcctcatc ttatattcat gtcttagctt gccaaccagg atatcagttt 5820
260 tccacataca tgtctctttc tgtcaagtag gatatcaagt tcccagggag gtcttggaaa 5880
261 cctaaacttt attctgcccc tactcaaaat ggaagtctta ttctaaatag gtacaggtgt 5940
262 ctctcttatg ttgggatcca tcccaagggc agcttaaaag gcaaatacta taaaggctga 6000
263 tacacaggtg cagaaagtgt tggtttctga gaacatccta gtaacagaag taacagcata 6060
264 tgagaaagtt tccttgtgat attaggaaca gcacaaactg gtaggttaga cgggtaacag 6120
265 ttaccaagac cttaacaatc ccagttcctc tttcaggtct tgagtggctg tattcctttc 6180
266 cttccattgt ttgtgctttc atagacttct ttaagggatt taatgttttc tttttaagga 6240
267 cctctagcat acacatatag gctgtgttaa ggtctttatg tgtgcttcca gggtgtaata 6300
268 ctcagggcct gctgtgatag ggttggtggg ttctagtgga gacctatcgt cctggctgta 6360
269 attggttgtg caggggtagc ctgtaggttc ccaatgagtg tgtgcctgag ctggatgctt 6420
270 gggaaaaaca ttgagtgacg ggaggaaagt ggggggccag ggatctgtat gcttcactga 6480
271 agatgggtgc agaagcagcc tagggctgag actgaggggt tccactctga gaagcagagg 6540
272 gagaggtgaa gatctgcagt tagcccacct gcgtccctgc ccagtgtggc ctgtgggttc 6600
273 ccagggagtg ccggctggag ttgggggtqg agggtaggac agggcaatga gtgggggaag 6660
274 ggaatttagg aggggaagat ctgtgggatc caccagcgat gaggtggctg tggtggaagc 6720
275 cgctgcagga gttagcgcag agctcaggat gaaactaggg attgggcgtg gaggaatgga 6780
276 gqqaqcqtqq aqqtcqcctc tccctctccc tqqataggta ggtcacccat ttgcttcccg 6840
277 tcagagagtg cctaagagag ttggaggctg ctttcctggt tgaagttgga tagaactttt 6900
278 caaactatat taatctgatg tgaaataagc acgtgaaagt gaacctccag cactgaatgt 6960
279 tggctatttt ctaccagcct cagttcacct atgaatggag actccaggct gcatcgtccc 7020
```

Input Set : F:\UTSC772US.APP.txt

Output Set: N:\CRF4\07142006\J534099A.raw

## Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

```
Seq#:3; N Pos. 265,1987,2010,2018,2032,2339,2355,2375,2383,2436,2955,2959
Seq#:3; N Pos. 2962,2998,4626,8255,18422,18460,18464,18484,18495,18508
Seq#:3; N Pos. 18520,18524,18560,18570,18571,18576,18591,18594,18610,18613
Seq#:3; N Pos. 18614,18619,18632,18633,18634,18635,18636,18637,18638,18639
Seq#:3; N Pos. 18640,18641,18642,18643,18644,18645,18646,18647,18648,18649
Seq#:3; N Pos. 18650,18651,18652,18653,18654,18655,18656,18657,18658,18659
Seq#:3; N Pos. 18660,18661,18662,18663,18664,18665,18666,18667,18668,18669
Seq#:3; N Pos. 18670,18671,18672,18673,18674,18675,18676,18677,18678,18679
Seq#:3; N Pos. 18680,18681,18682,18683,18684,18685,18686,18687,18688,18689
Seq#:3; N Pos. 18700,18701,18702,18703,18704,18705,18706,18707,18708,18709
Seq#:3; N Pos. 18710,18711,18712,18713,18714,18715,18723,18735,18855
```

DATE: 07/14/2006 VERIFICATION SUMMARY TIME: 10:18:52 PATENT APPLICATION: US/10/534,099A

Input Set : F:\UTSC772US.APP.txt

Output Set: N:\CRF4\07142006\J534099A.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number L:167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:240

M:341 Repeated in SeqNo=3